DOCUMENT RESUME

ED 221 573 TM 820 602

AUTHOR Haynes, Lathrop T.; Cole, Nancy S.
TITLE Testing Some Assumptions About On-Level Versus Out-of-Level Achievement Testing.
PUB DATE Mar 82
NOTE 20p.; Presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, March 20-22, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Tests; Elementary Secondary Education; Instructional Program Divisions; Motivation; *Student Motivation; Student Teacher Relationship; *Testing; Testing Problems; *Test Reliability
IDENTIFIERS *Out of Level Testing

ABSTRACT

The study was concerned with the merits of functional level versus nominal level achievement testing. The reliability of test scores when students were tested at their functional levels rather than their nominal levels was examined. Motivation and frustration of students was studied during testing at their functional levels. Another goal of the study was to examine the relationship between teacher judgments and locator test scores. A determination was made as to how each related to final test scores achieved. The 3R's Test, a general achievement test battery was used in the study. The results of the testing were discussed in terms of the three areas of concern, and generally confirmed the theoretical expectation that appropriate assignment to functional level produced more reliable scores for a larger proportion of students tested. Some theoretical and common sense assumptions about the merits of functional level testing were confirmed. However, the results in terms of proportion of in-range scores and student reports about the testing were of a smaller magnitude than expected. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L.T. Haynes

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED221573

Testing Some Assumptions About On-Level Versus Out-of-Level Achievement Testing

Lathrop T. Haynes and Nancy S. Cole
University of Pittsburgh

TM 820602

When a student is performing behind or ahead of his or her nominal grade level, one faces the decision of whether to test the student on content associated with the current nominal grade level or on content appropriate to the current level of performance. Testing with the level of the test designed for the current nominal grade level is typically called "on-level" or "in-level" testing. When the student's functional level is different from the nominal grade and the student is tested at the functional level rather than the nominal level, the term "out-of-level" testing is applied. Measurement theory dictates that the most efficient measurement occurs when the test is at the student's functional level. If a test is too hard, there may be increased guessing and frustration and decreased motivation. If a test is too easy, boredom and carelessness may also result in a poor measurement. Thus, theoretical expectations favor functional level testing--testing with a level of the test which is consonant with the student's performance level. For students functioning above or below current grade placement, functional level testing implies out-of-level testing.

When out-of-level testing is to be employed, the problem arises of how to determine the student's functional level. The methods for determining students' functional levels and assigning test battery levels to these students have typically been limited to the use of locator tests, teacher judgments, and the results of prior tests. These methods of assigning levels, however, may not all produce the same classification of students to levels of the test, and

which method actually does the best job in predicting students' actual functional levels is not known.

The main purpose of this study was to investigate the relative effectiveness of out-of-level versus on-level testing (1) in improving the reliability of the scores by reducing floor and ceiling effects, and (2) in improving motivation and reducing frustration. The study also examined the levels of agreement among two methods of assigning test battery levels to students and the levels of agreement of these methods with final test scores.

Statement of the Problem

The research questions to which this study is addressed are:

Research question 1: When students' functional levels are different from their nominal levels, does testing them at their functional levels produce more reliable scores than testing them at their nominal grade level?

Most studies of out-of-level testing approach the reliability issue by examining the proportion of scores lying within the range of reliable scores bounded by the floor or chance level as the lower limit and by the ceiling level as the upper limit (Ayre & McNamara, 1973; Crowder & Gallas, 1978; Cleland & Idstein, 1980; Long, Shaffran & Kellogg, 1977; Powers & Gallas, 1978; Yoshida, 1976). The theoretical expectation is, of course, that testing at functional levels below the nominal level produces more reliable results than on-level testing mainly because the on-level test is too difficult and results in more scores below the chance level. Likewise, it is assumed that testing at functional levels above the nominal level will produce more reliable results than on-level testing because the on-level test is too easy, resulting in scores grouped above the ceiling level of the test, a region in which carelessness may play an unduly large role in score differences and differentiation between

students may be inexact. "In-range" scores (scores between the floor and ceiling) are expected to be more reliable scores. Thus the question of reliability is addressed by considering the proportion of scores in-range for functional level and nonfunctional level testing.

Research question 2: Is frustration lessened and a better effort put forth by students tested at their functional levels rather than at their nominal grade levels?

Motivation and frustration are interrelated but different dynamics which affect the performance of students on tests. Students for whom a test is too difficult may be frustrated by the difficulty level and unmotivated. Students for whom a test is too easy may not be adequately motivated either and be bored and perhaps careless with the too-easy test. Students whose responses are coming from frustration or boredom may not be producing their maximum performance. When the appropriate level of a test is administered, frustration, low motivation, and the tendency to drop out of competition may be minimized resulting in a better measure of achievement (Coffman, 1979).

For students whose functional level is below the nominal level, it is typically assumed that frustration is lessened and motivation is improved by testing near the functional level. Less has been discussed about the effects of above-level testing, although there too, functional level testing may be expected to produce optimal motivation for students for whom the on-level test is too easy. Actual reactions of students to these different testing conditions have not been examined to see if the expected effects occur.

Research question 3: How do locator test scores and teacher ratings of the level at which students should be tested compare in predicting test scores produced by the achievement test battery?

The decision of how to determine what level of a test to administer when out-of-level testing is being considered is usually limited to three possible methods: 1) the results from previous testings, 2) teacher ratings, and 3) locator tests. Teachers' ratings likely reflect the results of previous tests (Hieronymous, 1977), and most discussions assume that teacher judgments should supplement any procedure used (Haenn & Proctor, 1978). Nevertheless, it has been shown that when teachers had access to locator test scores, they used them (Slaughter & Gallas, 1978, p. 3). With these relationships in mind, the goal was first to examine how closely related the two major assignment methods, teacher judgments and locator test scores, are, and second, to see how each relate to the test scores finally achieved.

Instrumentation

The objective test. The test battery used in this study is The 3R's Test, a general achievement test battery published by the Riverside Publishing Company (Cole, Trent, & Wadell, 1982). The battery consists of nine separate levels for grades kindergarten through grade 8 plus two combined levels for grades 9-10 and 11-12 and is marketed by the publisher as one battery with three sections, Reading, Language, and Mathematics. Each level is designed to be appropriate in content and difficulty to generally recognized curriculum at each grade.

All levels of The 3R's Test used here are four-option multiple choice, and above level 8, grade two, answers are recorded by the student on a separate machine-scorable answer sheet. Level 8, grade 2, does not contain a language section and answers are recorded in the machine scorable answer booklet.

The floor level for each subtest was defined as k/a, where k is the number of items on the test and a is the number of options per item, which is the expected score if a student were simply guessing (Gullikson, 1950). The ceiling

level was set at 90% correct for each section. The number of items per subtest and the raw score floor and ceiling levels for each section appear in Table 1.

The locator test. A series of locator tests is also available from Riverside for use with The 3R's Test. Since the battery is not divisible into three separate subject tests, the locator tests are designed to get a quick appraisal of students' achievement across all three subject areas combined. The locator tests used with this study were in their preliminary stage of development but were nonetheless expected to give reasonable rankings of student's overall achievement; however, at this stage, the locator test did not yield a specific test level assignment. Four locator tests are available, one each for use in grades K-3, 4-5, 6-7, and 8-12. The locator tests used in this study were the ones for grades 4-5 and for grades 6-7. Each locator test consisted of 20 four-option multiple-choice items. Answers were recorded by the student in the locator test booklet which was hand scored.

Based on preliminary data on the locator test supplied by the publisher, the raw score range for each locator test was divided into five sections intended to give adequate numbers of students in each category for the study and possibly to indicate a roughly appropriate test level for each student as well. At grade 4, the five locator test raw scores categories used were: (1) 5 and below; (2) 6 through 7; (3) 8 through 10; (4) 11 through 13; and (5) 14 and higher. At grade 6, the categories were: (1) 6 and below; (2) 7 through 8; (3) 9 through 11; (4) 12 through 13; and (5) 14 and higher.

Teacher rating form of student performance. The teacher's rating form used in this study consisted of a brief statement of the problem of testing students at one grade level when their schoolwork is actually more typical of another grade. Teachers were asked to examine a set of test booklets ranging from two levels below the class's level to two levels above the class's level

and to identify students who could reasonably be tested with a level other than their grade level test. Students who were not categorized into one of the four out-of-level categories were assumed to be performing typically for their grade level and could most appropriately be tested on-level.

Student questionnaire. The student questionnaire consisted of seven three- or four-option multiple choice questions which attempted to measure students' perceptions of the test battery. The questionnaire asked students whether they tried hard or gave up on the test, and whether they were frustrated or bored by taking the test. Students recorded the answers to the questionnaire on the questionnaire sheet as the teachers read the questionnaire aloud to their classrooms at the end of the testing session. The questionnaire took approximately three minutes to administer.

The Sample and Data Collection

The data collection took place in three phases: (1) the administration of the locator tests and collection of teachers' ratings, (2) the scoring the locator tests and assigning students to levels of the test battery, and (3) the administration of the test batteries and student questionnaires.

The schools selected to participate in this study were in eastern Ohio and were selected because they were believed to have had a relatively high proportion of students performing out of level (primarily below level). Four hundred seventy six fourth grade and 472 sixth grade students were tested.

Procedures

On the first day of testing the locator tests were administered in each school. Locator testing in each classroom took 15 to 20 minutes including instructions. While the students took the locator test, the teachers rated their students' achievement levels on the teacher's rating form.

The locator tests were then scored and students were grouped according to their locator test raw score into one of the five categories for each grade. Then, within each category, students were randomly assigned to take either an on-level or out-of-level test. Students in Category 1 took either the on-level test or the test two levels below grade level; students in Category 2 took either the on-level test or the test one level below grade level. Students in Category 4 took the on-level or the one-level-above-grade-level test; and students in Category 5 took the on-level or the two-levels-above-grade-level test. Students getting approximately average scores for the grade (Category 3) were randomly assigned to one of the five test levels for that grade. The design produced the numbers shown in Tables 2 and 3 for each combination of locator test category and level of The 3R's Test administered.

On the morning of the second day teachers administered The 3R's Test to their classes. Students taking both the out-of-level tests and on-level test were tested together in their own classrooms (except for those students taking Level 8, grade 2, who were tested separately). At the conclusion of the testing, each teacher read aloud the student questionnaire as the students read along silently and marked their answers.

Results

In-range scores for functional and on-level testing. Research question 1 asked whether functional level testing produces more reliable scores than nominal level testing. This question was addressed using the traditional approach of observing the proportion of students scoring below the chance level, above the ceiling level, and in between these two levels (in-range), for both the functional-level and on-level conditions. In part, the results were affected by the possible inaccuracy of assigning functional levels because of somewhat arbitrarily chosen cut scores for each functional level on the locator test.

Because of this, a larger number of fourth graders were categorized into functional levels below their nominal level (into Categories 1 & 2) and a larger number of sixth graders were categorized into functional levels above their nominal level (into Categories 4 & 5) than would be expected using locator cut scores which had been empirically tested. Nevertheless, the locator test scores did achieve a fairly reasonable assignment of students at both grades four and six.

The expectation for students in Categories 1 & 2 is that the proportion of students scoring below the chance level will be smaller for those students being tested below level (presumably nearer their functional level) than for those being tested on level. The expectation for the higher scoring students, those in Categories 4 & 5, is for the proportion of students scoring above the ceiling level to be smaller for students tested above level (presumably nearer their functional level) than for students tested at their nominal grade level.

The proportion of students scoring below chance level, in-range, and above ceiling level are shown in Table 4 and Table 5 for the fourth and sixth grades respectively. An "*" indicates the cells for which the test given is presumed to be nearer the student's functional level.

The results generally support the results of previous studies and the expectations of this study, although less dramatically than might have been expected. First, the proportions of students in the two low locator test categories, 1 and 2, who scored below the chance level on the on-level test were smaller than expected in both grades four and six. Nevertheless, the use of the below-level functional level test did reduce the proportion of students scoring below the chance level to zero in all cells for the Reading and Math subtests, but in some cases did not reduce or eliminate the students scoring below chance for the Language subtest. In some cases, especially in the fourth

grade, the proportion of students scoring above the ceiling level on the below-level test was greater than the proportion scoring below the chance level on the on-level test. This is likely due to the use of less than optimal locator test cut scores for the lower scoring groups, especially in Grade 4.

Second, the proportion of students with high locator test scores (those in Categories 4 & 5) who scored above the ceiling level was smaller in most cases, as expected, for those testing above level (presumably nearer their functional level) than for those testing on-level in both grades four and six. This was accomplished without increasing the below-chance-level scores for this group tested above level.

Student reports of effort and frustration. Research question 2 involved whether students actually report feelings of frustration and putting forth effort usually expected with functional level and nominal level testing. It was expected that students scoring low on the locator test (those in Categories 1 & 2) who were tested below level (presumably nearer their functional level) would report (1) feeling less frustrated, (2) trying harder, and (3) less often giving up than similarly low students tested on level. It was also expected that students scoring higher on the locator test (Categories 4 & 5) who were tested above level (presumably nearer their functional level) would report (1) trying harder and (2) less boredom than such students tested on level (not at their functional level. Middle scoring students (Category 3) tested on level (presumably near their functional level) were expected to report (1) less frustration, (2) trying harder, (3) less often giving up, and (4) less boredom than such students tested out of level (not at their functional level).

The results shown in Tables 6 and 7 for each questionnaire item tended to support the expectations. Perhaps of greatest interest is that although the

trends seem to follow the expectations, the magnitude of differences between cells on each question is generally not very large. This may be due in part to the few questionnaire item options providing little spread, but since it is consistent with relatively small effects of the functional level testing on the in-range scores, it may reasonably be interpreted to represent rather small perceived differences by the students.

Relation of locator scores, teacher ratings, and achieved scores. The third research question addressed how locator test scores and teacher ratings of the level at which students should be tested compare to each other and how each relates to the eventual test scores produced by the achievement test battery.

The correlations between teacher ratings, locator test scores, and achievement test composite scores appear in Tables 8 and 9 for grades four and six respectively. Correlations between teachers' ratings with locator test scores were .49 and .65 for grades four and six respectively, indicating that the two measures of assigning students to levels are generally related but certainly do not yield identical results.

Among all those students tested on level, correlations of teachers' ratings and locator test scores with The 3R's composite raw scores were very similar: .61 and .68 respectively for grade 4 and .72 and .76 for grade 6. Locator test scores correlated slightly higher than teacher ratings with composite raw scores which is not surprising since the content and type of questions on the locator test and the test battery are quite similar. If final achievement is used as a criterion for these two assignment methods, there seems to be no particular reason to prefer one over the other in terms of the level of correlation. For these variables, however, there is the issue of possible differences in the cutting points in assigning levels--an issue which is not addressed here.

Discussion

The results of this study generally confirm the theoretical expectation and previous results of others that appropriate assignment to functional level can produce a larger proportion of in-range scores and hence presumably more reliable scores for a larger proportion of the group tested. In this study, however, the locator score categories used to assign level for low scoring fourth graders were not optimal, and the problem of reaching the ceiling with too low a level test was apparent.

When students were asked about how hard they tried and whether they were bored or frustrated with the test, they tended to confirm the expected better motivation, less frustration, and less boredom for functional level as compared to nonfunctional level testing, but the differences were relatively small. Together these results tend to confirm some theoretical and common-sense assumptions about the merits of functional level testing, although the magnitude of the results in terms of proportion of in-range scores and student reports about the testing were perhaps of smaller magnitude than expected.

References

Ayre , J. E., & McNamara, T. C. Survey testing on an out-of-level basis. Journal of Educational Measurement, 1973, 10, 79-84.

Cleland, W. E., & Idstein, P. M. In-level versus Out-of-level testing of sixth grade special education students. Paper presented at the Annual Meeting of the National Council on Measurement in Education (Boston, MA, 1980). (ERIC Document Reproduction Service No. ED 187 725)

Coffman, W. E. Classical test development solutions. Number 23. Iowa Testing Programs Occasional Papers. Iowa City: Iowa Testing Programs, 1979. (ERIC Document Reproduction Service No. ED 175 913)

Cole, N. S., Trent, E. R., & Wadell. D. C. The 3R's Test, Achievement Edition (Grades K-12). Chicago: The Riverside Publishing Company, 1982.

Crowder, C. R., & Gallas, E. J. Relation of out-of-level testing to ceiling and floor effects on third and fifth grade students. Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, 1978). (ERIC Document Reproduction Service No. ED 155 212)

Gulliksen, H. Theory of Mental Tests. New York: John Wiley & Sons, Inc., 1950.

Haenn, J. F., & Proctor, D. C. A practitioner's guide to out-of-level testing. Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, 1978). (ERIC Document Reproduction Service No. ED 171 728)

Hieronymus, A. N. Individualized testing with the Iowa Tests of Basic Skills. In Putting Current Ideas in Measurement to Work in the Schools. Proceedings from the Thirteenth Annual Houghton Mifflin Measurement Conference, Iowa City, Iowa, 1977.

Long, J. V., Schaffran, J. A., & Kellogg, T. M. Effects of out-of-level survey testing on reading achievement scores of Title I, ESEA students. Journal of Educational Measurement, 1977, 14, 203-213.

Powers, S., & Gallas, E. J. Implications of out-of-level testing for ESEA Title I students. Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, 1978). (ERIC Document Reproduction Service No. ED 175 930)

Slaughter, H. B., & Gallas, E. J. Will out-of-level norm-referenced testing improve the selection of program participants and the diagnosis of reading comprehension in ESEA Title I programs?. Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, 1978). (ERIC Document Reproduction Service No. ED 155 213)

Yoshida, R. K. Out-of-level testing of special education students with a standardized achievement battery. Journal of Educational Measurement, 1976, 13, 215-221.

Table 1

Total Number of Items and Floor and Ceiling Levels by Subtest and Level of The 3R's Test

		Reading			Language			Math		
Level	Grade	No. Items	Floor*	Ceiling*	No. Items	Floor*	Ceiling*	No. Items	Floor*	Ceiling*
8	2	35	8	32	---	---	---	30	7	27
9	3	50	12	45	40	10	36	35	8	32
10	4	50	12	45	40	10	36	35	8	32
11	5	50	12	45	40	10	36	35	8	32
12	6	50	12	45	40	10	36	35	8	32
13	7	50	12	45	40	10	36	35	8	32
14	8	50	12	45	40	10	36	35	8	32

*The numbers shown describe the inclusive upper limit of the floor region and the inclusive lower limit of the ceiling region. In-range scores are those greater than the floor upper limit and less than the ceiling lower limit.

Table 2

Number of Students at Each Locator Test Level and Tested at Each of the 3R's Test Levels: Grade 4

3R's Test Level	Locator Test Score Category 1	2	3	4	5
Level 8 (Grade 2)	54		35		
Level 9 (Grade 3)		60	29		
Level 10 (Grade 4)	53	58	24	43	10
Level 11 (Grade 5)			29	40	
Level 12 (Grade 6)			31		10

Table 3

Number of Students at Each Locator Test Level and Tested at Each of the 3R's Test Levels: Grade 6

3R's Test Level	Locator Test Score Category 1	2	3	4	5
Level 10 (Grade 4)	26		24		
Level 11 (Grade 5)		33	22		
Level 12 (Grade 6)	36	31	25	38	76
Level 13 (Grade 7)			24	37	
Level 14 (Grade 8)			23		77

Table 4

Proportions of Students Scoring Below Chance (BC) Level, In-Range (IR), and Above Ceiling (AC) Levels by Subtest and Test Level, Grade 4

Locator Test Category	Level Administered		Reading			Language			Math		
			BC	IR	AC	BC	IR	AC	BC	IR	AC
1	2 Below Level	*		.67	.33	---	---	---		.80	.20
	On Level		.13	.87		.03	.89	.04	.08	.92	
2	1 Below Level	*		.85	.15	.07	.85	.08		.97	.03
	On Level		.05	.95		.03	.95	.02	.09	.91	
3	2 Below Level			.20	.80	---	---	---		.29	.71
	1 Below Level		.04	.64	.32	.04	.75	.21		.93	.07
	On Level	*		.96	.04		.96	.04		1.00	
	1 Above Level			1.00		.07	.93			1.00	
	2 Above Level			1.00		.10	.90		.16	.84	
4	On Level			.72	.28	.03	.63	.35	.03	.98	
	1 Above Level	*		.90	.10	.03	.93	.05	.03	.98	
5	On Level			.60	.40		.50	.50		.90	.10
	2 Above Level	*		.90	.10		1.00			1.00	

*Level estimated closest to functional level.

Table 5

Proportions of Students Scoring Below Chance (BC) Level, In-Range (IR), and Above Ceiling (AC) Levels by Subtest and Test Level, Grade 6

Locator Test Category	Level Administered		Reading			Language			Math		
			BC	IR	AC	BC	IR	AC	BC	IR	AC
1	2 Below Level	*		1.00		.08	.85	.08		1.00	
	On Level		.06	.94		.03	.92		.03	.97	
2	1 Below Level	*		.97	.03		1.00			.97	.03
	On Level			1.00		.03	.94	.03		1.00	
3	2 Below Level			.83	.17		.71	.29		1.00	
	1 Below Level			.91	.09	.05	.95			1.00	
	On Level	*		.92	.08	.04	.92	.04		1.00	
	1 Above Level		1.00			.04	.96			1.00	
	2 Above Level		1.00				1.00		.04	.96	
4	On Level		.03	.97			.95	.05		1.00	
	1 Above Level	*		.97	.03		.97	.03		1.00	
5	On Level			.74	.26	.01	.75	.24		.96	.04
	2 Above Level	*	.01	.97	.01		.96	.04		1.00	

*Level estimated closest to functional level.

Table 6

Means and Standard Deviations of Questionnaire Items by Test Level Taken, Grade 4

Category	Level Administrated		Did you try hard on this test? 4=Yes, on all. 3=Yes, on most. 2=No, not very hard. 1=No, not at all	Were you frustrated or upset because the test was too hard? 3=Yes, a lot. 2=Yes, a little. 1=No	Did you feel bored and impatient because the test was too easy? 3=Yes, a lot. 2=Yes, a little. 1=No	Did you give up because the test was too hard? 3=Yes, on much. 2=Yes, on a little. 1=No
1	2 Below Level	*	3.889 .372	1.173 .550	1.870 .825	1.037 .191
	On Level		3.725 .568	1.724 .777	1.549 .730	1.255 .483
2	1 Below Level	*	3.883 .454	1.525 .679	1.550 .769	1.100 .354
	On Level		3.793 .487	1.724 .720	1.293 .593	1.241 .432
Locator Test Category 3	2 Below Level		3.971 .169	1.029 .169	1.771 .843	1.000 .000
	1 Below Level		3.793 .412	1.379 .561	1.345 .614	1.103 .310
	On Level	*	3.833 .381	1.625 .576	1.375 .576	1.333 .482
	1 Above Level		3.828 .384	1.724 .649	1.310 .541	1.207 .412
	2 Above Level		3.759 .435	1.767 .626	1.600 .724	1.300 .466
4	On Level		3.825 .385	1.575 .594	1.300 .516	1.275 .506
	1 Above Level	*	3.769 .485	1.641 .668	1.421 .683	1.154 .366
5	On Level		3.900 .316	1.400 .516	1.700 .675	1.100 .316
	2 Above Level	*	4.000 .000	1.500 .527	1.500 .707	1.100 .316

*Level estimated closest to functional level.

Table 7

Means and Standard Deviations of Questionnaire Items by Test Level Taken, Grade 6

Category	Level Administrated		Did you try hard on this test? 4=Yes, on all. 3=Yes, on most. 2=No, not very hard. 1=No, not at all	Were you frustrated or upset because the test was too hard? 3=Yes, a lot. 2=Yes, a little. 1=No	Did you feel bored and impatient because the test was too easy? 3=Yes, a lot. 2=Yes, a little. 1=No	Did you give up because the test was too hard? 3=Yes, on much. 2=Yes, on a little. 1=No
1	2 Below Level	*	3.846 .368	1.358 .647	1.423 .504	1.192 .491
	On Level		3.583 .554	1.686 .676	1.389 .599	1.222 .485
2	1 Below Level	*	3.576 .502	1.455 .506	1.485 .667	1.455 .564
	On Level		3.645 .551	1.516 .508	1.484 .677	1.194 .402
Locator Test Category 3	2 Below Level		3.792 .415	1.292 .550	1.583 .717	1.125 .338
	1 Below Level		3.364 .790	1.545 .671	1.318 .477	1.455 .671
	On Level	*	3.680 .476	1.400 .500	1.240 .523	1.400 .500
	1 Above Level		3.625 .495	1.542 .588	1.292 .550	1.333 .482
	2 Above Level		3.652 .573	1.652 .775	1.217 .518	1.478 .665
4	On Level		3.632 .489	1.447 .504	1.316 .525	1.184 .394
	1 Above Level	*	3.703 .463	1.622 .639	1.405 .599	1.216 .417
5	On Level		3.697 .517	1.224 .419	1.355 .534	1.237 .428
	2 Above Level	*	3.592 .546	1.553 .641	1.132 .340	1.329 .500

*Level estimated closest to functional level.

Table 8

Pearson Product-Moment Correlation Coefficients Among Teacher Ratings (TR), Locator Test/Scores (LSR), and Composite Test Scores (COMP)
Grade 4

	TR	LSR
LSR	.49	
COMP	.61	.68

Table 9

Pearson Product-Moment Correlation Coefficients Among Teacher Ratings (TR), Locator Test/Scores (LSR), and Composite Test Scores (COMP)
Grade 6

	TR	LSR
LSR	.65	
COMP	.72	.76